DOCUMENT RESUME

ED 065 505 24 TM 000 918

AUTHOR Sullivan, Howard J.
TITLE Improving Learner Achievement through Evaluation by Objectives.
INSTITUTION Southwest Regional Educational Lab., Inglewood, Calif.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO SR-4
BUREAU NO BR-6-2865
PUB DATE 29 May 68
NOTE 29p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Academic Achievement; Behavioral Objectives; Behavior Change; *Classification; Course Content; Curriculum Design; *Educational Objectives; Educational Programs; *Evaluation Techniques; Improvement; Individual Development; Learning Activities; Skill Development; Standards; *Student Behavior; Task Performance; Taxonomy; Thought Processes
IDENTIFIERS Behavioral Classification System; Bloom Taxonomy

ABSTRACT

Evaluation techniques were designed to improve learner performance through use of pre-specified popular instructional objectives. Current curriculum planning and evaluation practices are examined. Two common evaluation malpractices are: (1) the tendency to treat the content of the program as the most important criterion for evaluation, (2) the evaluation of desired learner behaviors that occur under certain classroom conditions, but do not represent actual behavior changes. Two steps to aid in the development of desired learner capabilities are: (1) To list those capabilities that the learner should acquire from the course; and (2) To attempt to develop these learner capabilities through appropriate instructional experiences. The use of a standard classification system for specifying intended educational outcomes and classifying student behaviors is a procedure that has gained considerable acceptance. One such system is the Bloom Taxonomy, which classifies educational objectives "related to mental acts or thinking" into six categories: knowledge, comprehension, application, analysis, synthesis, and evaluation. Another system is a Behavioral Classification System published by the American Association for the Advancement of Science. Instructional objectives are important for two purposes: planning instruction and assessing its effects. Two types of evaluation are formative and summative. The former evaluates programs still under development and the latter completed programs. A need exists for further study of evaluation procedures. (CK)

DE-NCERD
TM
BR-6-2865
PA24

SWRL

SOUTHWEST REGIONAL LABORATORY FOR EDUCATIONAL RESEARCH & DEVELOPMENT

ED 065505

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

SWRL

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

TM 000 918

Improving Learner Achievement Through Evaluation By Objectives

SR 4 29 May 1968

© 1970 SOUTHWEST REGIONAL LABORATORY for Educational Research and Development. [illegible]

Published by Southwest Regional Laboratory for Educational Research and Development, pursuant to a contract with the National Center for Educational Communication, [illegible] tory by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the Office of Education, and no official endorsement by the Office of Education should be inferred.

ED 065505

IMPROVING LEARNER ACHIEVEMENT THROUGH EVALUATION BY OBJECTIVES

Howard J. Sullivan

In recent years the quality of education in American schools has received close scrutiny from both the general citizenry and the federal government. Massive federal programs have been enacted in an attempt to upgrade student achievement in the schools. One potentially healthy aspect of these programs is the requirement that evaluation of the effects of federally supported school projects be conducted as a part of the project. This federal legislation has been responsible for a new emphasis on educational evaluation.

Extensive evaluation in schools cannot be justified unless its ultimate consequence is the improvement of education. Good evaluation procedures should lead to improved student achievement. This paper, therefore, will outline certain evaluation techniques designed to improve learner performance through the use of pre-specified instructional objectives. Several curriculum planning and evaluation practices that enjoy current popularity will also be examined.

This paper was prepared by the author as a member of the American Educational Research Association Division B Committee on the Stating of Instructional Objectives. It is one of four papers written by individual committee members on the topic of educational objectives.

The overriding concern in educational evaluation is the behavior and capabilities acquired by the child. The child enters school lacking certain abilities and behaviors. Educational experiences are provided so that he will acquire these capabilities. When a child consistently performs a given act in a situation in which he previously did not perform it, he demonstrates that he has acquired a particular capability. The essential criterion for evaluating educational experiences is the extent to which they produce desired learner behaviors that previously were not performed under the same conditions. These changes in learner behavior are the outcomes with which educational evaluation must be most concerned.

EVALUATION WITHOUT POST-INSTRUCTIONAL OBJECTIVES

At first glance it appears elementary to stress that changes in learner performance are the most appropriate single source for evaluating educational methods and materials. Yet, the success of teachers and educational programs is often judged primarily or exclusively on other criteria. Two evaluation malpractices that employ criteria other than post-instructional learner outcomes as the principal source of evaluation currently occur with sufficient frequency to be worthy of mention.

One common error in the evaluation of school programs results from the tendency of many educators to treat the content of the program as the most important criterion for evaluation. The content of an educational program is simply the materials and methods employed by the teacher. Barring public relations complications and financial considerations, any given content may be installed at will by a teacher or school administrator. Although it should be apparent that the content of a program is an educational means, not a goal in itself, it is not always treated as such. From time to time certain methods and materials, often classified as "innovative," become cherished in educational circles. For example, the present educational climate is such that few self-respecting teachers would admit that they do not employ a problem-solving or discovery-learning approach. The popularity of these programs and program components is often based more upon some sort of intrinsic appeal or other elusive factors than upon empirical evidence of their effectiveness. Nevertheless, teachers and educational programs are often evaluated on the basis of whether or not they employ certain favored methods and types of materials, and little attempt is made to determine the effectiveness of this content in improving learner performance. The presence or absence of discovery procedures, individualized instruction, multi-media materials, or a multi-sensory approach clearly is not an appropriate criterion for evaluating instruction.

A second inappropriate basis for evaluation involves desired learner behaviors that occur under certain classroom conditions but do not represent actual behavior changes in the learner. Student enjoyment, involvement, and self-expression in the classroom are conditions that imply certain pupil behaviors whose frequent occurrence in class is highly desired by

teachers. "Fun-type" classroom activities, often containing related instruction, are introduced in the hope that students will enjoy them and participate actively in them. The activities selected are such that the student enjoyment and participation would occur in response to the same types of activities at either an earlier or a later date without the particular enjoyable classroom sequence. Rarely is any evidence obtained that the children will subsequently be happier, more involved, or more self-expressive in either the same or different situations. In short, the presence of these conditions in the classroom normally does not constitute student learning. They are responses that the child has already learned to make in similar situations, and the classroom activity provides him with an additional opportunity to make these learned responses. Jackson (1966) records circumstances of this type when he notes with approval that

> More and more I have come to think of the teacher's work as consisting primarily of making some kind of an educated guess about what would be a beneficial activity for a student or group of students and then doing whatever is necessary to see that the participants remain involved in that activity. The teacher's goal, in other words, is student involvement rather than student learning.

Unfortunately, this condition appears to be true. It is so reinforcing to see students happily involved in a classroom activity that it is tempting to conclude that the instruction must be highly successful. Yet, with the many desirable instructional outcomes that are essential to the learner's "happy involvement" and success in later life, it seems shallow to consider as important sources of evaluation in-class behaviors or conditions which do not represent new learning for the child. The student who spends a happy, involved, self-expressive educational career in the classroom but fails to acquire basic reading and mathematics skills (or appropriate inter-personal behaviors, if you prefer) is a sad product of an educational system.

The present intent certainly is not to derogate student enjoyment, involvement, or self-expression. Far from it! It would be most worthwhile if one could identify the behavioral components of these conditions and provide classroom experiences that result in a subsequent increase in their occurrence in important situations. Also, the presence of such conditions may often be a contributing factor to student attainment of desired instructional outcomes. The matter of concern at this point is simply that the mere in-class occurrence of these conditions in response to classroom activities selected to elicit them is not an appropriate criterion for evaluating instruction. Here, the writer shares the feeling expressed by Stanley (1967) in discussing the school experience of young children:

> I feel quite uneasy lest expressing oneself becomes in itself the main goal, rather than a means toward later goals of a more educational-vocational sort.

Our schools are charged with the responsibility for developing desirable post-instructional behaviors that the learner previously does not perform in given situations. The criterion of utmost concern in evaluating instruction is the presence or absence of these learner behaviors in appropriate situations in which the behaviors previously did not occur.

OBJECTIVES AND CURRICULUM PLANNING

The idea that children should learn things in school is neither revolutionary nor controversial. Current evaluation malpractices notwithstanding, few educators deny that certain capabilities should be acquired by the learner as he makes his way through school. However, the method by which these capabilities should be developed and are developed is a matter of considerable controversy among educators.

Two logical steps to take in attempting to develop learner capabilities that are desired outcomes of a particular course of instruction are (1) to list those capabilities that the learner should acquire from the course and (2) to attempt to develop these learner capabilities through appropriate instructional experiences. This approach has been used quite extensively in programmed instruction. In addition, many curriculum specialists express strong verbal commitment to such a procedure, but its actual use in the schools has been very limited when one considers its professed popularity among curriculum experts.

Vociferous opposition exists to this logical approach. Some educators consider it more democratic and humanistic to simply introduce content or let the children select it and see what behavior changes, if any, occur. The systematic design of instruction to achieve desired behavior changes is derogated as "mechanistic procedures" by its opponents. We are told by Arnstine (1964), for example, that the notion of pre-planned educational objectives "supports the unrestrained manipulation of human beings." Arnstine is, in fact, undoubtedly correct, although his statement hardly constitutes a valid argument against pre-specified objectives. What he fails to note is that the unrestrained manipulation of human beings is also supported by mandatory school attendance requirements, recess and lunch periods, stop lights, and school safety patrols, among other things.

The most frequent alternative to instructional planning offered by opponents of pre-planned objectives and instructional sequences is the selection of content by the teacher and/or the students and the presentation of that content in an unspecified, open-ended manner. The different emphasis and the instructional planning embodied in this approach is perhaps captured best in the oft-quoted statement of one of its leading proponents who states (Macdonald, 1965) that

> In the final analysis, it could be argued, the teacher in actuality asks a fundamentally different question from "What am I trying to accomplish?" The teacher asks "What am I going to do?" -- and out of the doing comes accomplishment.

Somehow, this "What-am-I-going-to-do" approach hardly seems more humanistic or democratic than one which attempts to develop pre-specified capabilities in the child. Both logic and the limited available research evidence (reviewed by Anderson, 1967) support the pre-specification of desired outcomes in terms of learner behavior and the planning of instructional activities that are designed to achieve these outcomes. The use of this approach is clearly more consistent with the position that schools are responsible for developing certain capabilities in the child and not merely for the presentation of content.

The movement toward the statement of instructional objectives in terms of desired post-instructional behaviors to be performed by the learner has received great impetus in recent years. Two widely-claimed advantages for the statement and use of behavioral objectives in curriculum and instruction are (1) they enable the teacher to know exactly what behaviors the learner should be able to perform as a result of instruction, and consequently facilitate the selection of materials and activities to develop these behaviors, and (2) they permit valid assessment of whether or not students have acquired desired post-instructional behaviors, and thereby also indicate the effectiveness of the instruction. Techniques for capitalizing upon the claimed advantages of precise objectives in instruction are discussed in a recent paper by Popham (1968). The potential benefits of their use in evaluation are treated later in this paper.

Certain other suggested uses or concomitants of explicit instructional objectives deserve passing mention. Gagne (1965) has proposed that the attainment of desired learner outcomes may be increased by telling students the objectives of a unit prior to instruction. Mager (1962) notes that this practice enables the student who is studying for a particular course to select activities that are relevant to the teacher's goals. It is also probable that the use of precise objectives in planning instruction results in more classroom practice on relevant learning tasks. When the teacher's main concern is the development of specific learner behaviors, rather than the presentation of content, he is more likely to provide opportunities for practice of these behaviors during instruction. Several instructional episodes recorded in elementary school classrooms and transcribed by Sullivan, Baker and Schutz (1966) illustrate the inefficient instruction that can occur when the teacher presents content, but does not provide for learner practice of the desired behavioral outcomes.

Obviously, the statement of instructional objectives for a course or unit of instruction is of no use in itself. The stated objectives must serve as the referent for planning instruction that leads to their attainment and for evaluating the success of the instruction and of individual learners. The contrast between this type of instructional

planning and an approach that is often used is highlighted in Stanley's (1967) discussion of Project Head Start:

> Someone said, "Let's have six or eight weeks of Project Head Start and see what improvements seem to result from this period." Few have yet asked, "What does it take to accomplish certain specific goals."

There is a definite need for more empirical research data on various procedures in curriculum planning and development. In the meantime, curriculum planning that is based upon specified desired learner outcomes is certainly a more rational approach than simply selecting treatments with no particular objectives in mind for the learner.

OBJECTIVES AND TAXONOMIES

The use of a standard classification system for specifying intended educational outcomes and classifying actual or inferred student behaviors is a procedure that has gained considerable acceptance. It is generally believed that systems of this type assist in the process of curriculum development by suggesting desirable student behaviors that can be performed on given units of content. An added attraction of such classification systems is the fact that they furnish standard terminology to describe various learner performances, thereby enabling educators to communicate in terms with agreed-upon definitions.

THE BLOOM TAXONOMY

In terms of the attention and use that it has received, the signal work among extant classification systems is the Taxonomy of Educational Objectives, Handbook I: Cognitive Domain (hereafter referred to simply as "the taxonomy"), developed by Bloom (1956) and his colleagues. The taxonomy classifies educational objectives "related to mental acts or thinking" into six categories: knowledge, comprehension, application, analysis, synthesis and evaluation. Each of the six categories represents a mental process presumably performed by the student in a given instructional or test situation. The particular mental process is inferred from the situation plus what is known of the student's prior educational experiences, and there is evidence (McGuire, 1963; Stanley & Bolton, 1957) of high agreement among trained raters in classifying test items into the six categories. The six categories are ordered as listed above to represent a hierarchy such that each successive process is purportedly more complex than the preceding one and is built upon it plus all other preceding ones. A taxonomy of objectives in the affective domain (Krathwohl, Bloom & Masia, 1964) supplements the cognitive taxonomy.

The Bloom taxonomy has achieved widespread popularity among educators

and has served several functions. Its influence in stimulating interest in educational objectives is evident from the many taxonomy-based research studies summarized by Cox and Unks (1967). Teachers have often been chastised for requiring a high proportion of rote-learning responses from their students, and the taxonomy has been widely used in teacher-education endeavors in an attempt to train teachers and aspiring teachers to identify other types of appropriate student behaviors. In addition, it may well be that the taxonomy has also had an effect in shifting the instructional efforts of some teachers away from the presentation of content and toward operations performed by the learner.

Despite the popularity of the taxonomy and its success in stimulating interest in educational objectives, its usefulness as a tool for effective curriculum planning and development is quite limited. Any attempt to use the taxonomy in the formulation of objectives must take into account its lack of precision in indicating either specific overt behaviors to be performed by the learner or the conditions under which they will be performed. In addition to the content on which learner behavior is to be performed, a useful instructional objective must state both the intended observable learner behavior that will result from instruction and the relevant conditions under which it will be performed. Rather than identifying classes of observable learner behaviors that can be used in task description and task analysis, the categories of the taxonomy describe mental processes that are inferred from skills and capabilities described in general terms in the Bloom Handbook. The majority of the sample objectives listed in the Handbook are not stated in terms of observable learner behaviors. They also ignore the important question of the conditions under which the learner performance is expected to occur, except to the extent that this is implicit in some of the examples. If instructional objectives are to be useful in the sequencing and evaluation of instruction, they must be stated much more precisely than are the categories and objectives in the taxonomy. Thus, Krathwohl's (1964) statement that curriculum analysis using the taxonomy

> aids in placing the material in the program sequence and in planning the over-all development of the skill of ability

simply is not correct. The taxonomy's lack of specificity in dealing with task analysis and task description renders it useless for the purpose of sequencing instruction. At best, the taxonomy serves as a guide for describing very general desired outcomes of educational programs and for suggesting objectives which then must be stated in terms of observable learner behavior to be useful for evaluation and instructional purposes.

Perhaps the most serious problems with the taxonomy are related to the lack of evidence that there is any generalizability of the imputed mental processes across subject-matter content. The important question here is, "Does the student perform the same mental operations across the range of subject-matter content in education?" Bloom and his colleagues state (1956, p. 12) that it is their assumption that he does. In an extensive study, Kropp and Stoker (1966) were unable to reach a conclusion

about the generalizability of the six inferred processes across content. The content materials used by Kropp and Stoker were two expository articles in science and two in social science. Even though these researchers found no conclusive evidence, some empirical support for the generalizability of imputed mental processes over content would not be too surprising in the related fields of science and social science where the apparent organization of the subject-matter content (especially that presented by Kropp and Stoker) is quite similar. Suppose, however, that they had conducted a study of the inferred mental operations performed across four such subject areas as reading, social science, spelling and arithmetic, where both the tasks required of the student and the structure of the content differ sharply from subject to subject. If evidence was lacking for the transcendance of the six processes across similar science and social science content, what is the probability that generalizability of the imputed mental processes occurs across four areas that are as diverse as those mentioned above?

Unlikely as it may be, let us assume for the sake of argument that there is generalizability of the imputed processes over content. Another consideration immediately confronts us. Will training or practice with given mental operations on content within one subject-matter area increase the learner's ability to perform those operations in other areas or even on other content in the same area? Several characteristics of the taxonomy would make worthwhile research on this question virtually impossible, but the general research evidence on transfer of training leads one to conclude that it would be extremely naive to expect such training or practice to increase the learner's abilities with other content.

If training students to perform the six imputed mental processes does not increase their ability to perform them with other content, it makes little sense to emphasize these operations in curriculum planning or test construction. A better approach is to identify the important overt behaviors that the learner should acquire from the course, to employ instructional activities designed to develop the behaviors, and to examine the learner's performance to determine whether or not he has acquired them. The behaviors should be specified in terms of the important tasks that students should learn to perform with the given subject-matter content. It is not sensible to construct a formula that will ensure the inclusion of a certain proportion of various types of items or practice, irrespective of the desired learner outcomes for that particular content. For example, the curriculum planner formulating objectives for a reading curriculum should ask such questions as "What reading behaviors should the student acquire?" and "How can one tell when a student has acquired them?" The question, "How can sufficient opportunities be provided for analysis, synthesis, and evaluation?" may distract him from the more important issues and lead to an emphasis on less appropriate content and learner outcomes.

Several interesting observations and research findings involving the taxonomy relate to the above discussion. Wolf (1967) notes that

> The Taxonomy of Educational Objectives sets forth a description of cognitive processes while taxonomy-type tests are designed to evoke those processes in the solution of problems and the answering of questions. To what extent taxonomy-type tests do, in fact, evoke the intended processes has yet to be demonstrated.

Kropp and Stoker (1966) present evidence from their study with science and social science content that generally supports the imputed hierarchy of the taxonomy in these areas and indicates that test-item difficulty increases at each higher level. An effect of the current emphasis on "higher mental processes" is seen in the research of Bialek (1967), who reports the paradoxical finding that the number of teachers who classify the application of knowledge (level 3 in the taxonomic hierarchy) as essential is significantly greater than the number classifying the acquisition of knowledge (level 1) as essential. The implication here that knowledge can be applied without being acquired suggests the need for teacher training in the sequencing of instruction. Despite the professed preference of teachers for higher-level objectives, however, studies of questions asked in class and in teacher-made tests (e.g., Davis & Tinsley, 1967; Scannell & Stellwagon, 1960) reveal that well over 50 per cent are rote-memory types of items. In one such study (Lawrence, 1963), 98 per cent of the social studies test items collected from 63 high schools were classified at the knowledge level of the taxonomy.

It is entirely possible to formulate objectives for a particular unit of content simply by listing, without reference to a taxonomy or some other classification scheme, the behaviors that learners should acquire relative to that content. Yet, a classification system listing those learner behaviors that represent desired outcomes of instruction may frequently be useful to the teacher or curriculum designer in identifying important learner behaviors to be acquired. With a given unit of content, the categories in the classification scheme could be used as a checklist to identify desired en route and terminal learner behaviors relative to the content. Bloom (1956, p. 12) reports that some individuals have found his taxonomy useful in "seeing a possible range of objectives." However, a classification system for overt learner behaviors should be much more useful than the Bloom taxonomy in identifying appropriate behavioral outcomes of instruction, since it would deal directly with these observable outcomes rather than with mental processes that are inferred from them.

A BEHAVIORAL CLASSIFICATION SYSTEM

The list of ten performance descriptions published by the American Association for the Advancement of Science (1965), constitutes a classification system that is based upon overt learner behavior, and not upon inferred mental operations. As such, it contrasts rather sharply with the Bloom classification scheme. The AAAS list was originally developed

for use with the AAAS science series entitled Science -- A Process Approach. Nonetheless, the performance descriptions in this list are equally appropriate for describing learner outcomes in subject-matter areas other than science.

One major error was committed in constructing the ten-category AAAS list. In many instances considerable overlap exists because the same learner behavior may be classified into any one of several different categories, depending upon the content of an item or the past learning history of the student. Thus, as in the case of the Bloom taxonomy, many tasks or learner performances cannot be correctly classified solely on the basis of the characteristics of the task or performance. The observer must also be able to describe the relevant past educational history of the learner, and the same performance is classified in one category for learners with one set of educational experiences and in another category for learners with different previous experiences. This is both an undesirable and an unnecessary feature of a behavioral classification system.

A more functional classification scheme for describing and categorizing student behaviors can be developed by reducing the number of categories in the AAAS list and modifying the definitions of the remaining categories. It is possible with a total of six performance terms to classify nearly all of the learner behaviors related to cognitive tasks in school learning. These six performance terms and their definitions are presented below. Two sample objectives are listed for each performance term, and common equivalent terms that are subsumed under each category are given for that category. Gerlach and Sullivan (1967) describe this classification scheme in considerable detail and present numerous examples of objectives and test items for each category.

PERFORMANCE TERMS

IDENTIFY: The learner indicates membership or non-membership of specified objects or events in a class when the name of the class is given.

Examples:

1. Objective: Given a list of instructional objectives, the learner will identify those that are stated in terms of learner behavior.
   Test Item: Mark an X by the number of each behavioral objective in the following list.
2. Objective: Given examples of equilateral, isosceles and obtuse angles, the learner will identify the isosceles triangles. (He could also be asked, of course, to identify the examples of each type.)
   Test Item: Draw a circle around each isosceles triangle.

The learner may be asked to respond to each "identifying" item by underlining, circling, or otherwise marking his response choice, or by pointing, touching, speaking, etc.

Equivalent terms and phrases subsumed under the term "identify": select, distinguish between, discriminate between, mark, match.

NAME: The learner supplies the correct verbal label (in speech or writing) for a referent or set of referents when the name of the referent is not given.

Examples:

1. Objective: The learner will name the principles of learning illustrated, when shown previously unencountered filmed classroom episodes illustrating selected principles of learning.
   Test Item: Write the name of the principle of learning that is best illustrated by each episode.
2. Objective: Shown examples of isosceles triangles, the learner will name the type of triangle shown.
   Test Item: What is this type of triangle called?

Equivalent terms subsumed: label, list.

DESCRIBE: The learner reports the necessary categories of object properties, events, event properties and/or relationships relevant to a designated referent. The teacher should decide in advance the learner responses that will serve as acceptable descriptions, although he should also accept other given descriptions that he deems correct but did not anticipate.

Examples:

1. Objective: The learner will describe at least three major effects of specified changes in climate or topography on economic conditions in the western United States.
   Test Item: Describe three or more ways in which the economy of California would be affected by a permanent, statewide 15-degree decrease in average daily temperature combined with a 10-inch statewide increase in annual rainfall.
2. Objective: The learner will describe the characteristics of an isosceles triangle.
   Test Item: What is an isosceles triangle?

Equivalent terms and phrases: define, tell how, tell what happens when.

CONSTRUCT: The learner produces a product (e.g., a drawing, article of clothing or furniture, map, essay, examples of a particular concept, etc.) which meets specifications given either in class or in the test item itself.

Examples:

1. Objective: The learner will construct at least two statements of observation and two statements of inference about selected environmental phenomena.
   Test Item: Write two or more statements of observation and two or more statements of inference about the experiment that was just completed.
2. Objective: Given a ruler and compass, the learner will construct a triangle with two of its sides differing in length by no more than 1/8 inch.
   Test Item: Draw an isosceles triangle.

Equivalent terms: prepare, draw, make, build.

ORDER: The learner arranges two or more referents in a specified order. The learner may be required to name or describe the referents in order himself, or a group of referents may be provided for him to order.

Examples:

1. Objective: Given a list of 5-10 events from the "Golden Age of England", with each event having been either a cause or consequence of at least one other listed event, the learner will order them chronologically.
   Test Item: Indicate the order in which the following events occurred by numbering them from 1 to 7.
2. Objective: The learner will describe in order the steps in constructing an isosceles triangle. (Here, the learner is expected to describe and order.)
   Test Item: Describe in order the steps involved in drawing an isosceles triangle.

Equivalent terms and phrases: arrange in order, sequence, list in order.

DEMONSTRATE: The learner performs the behaviors essential to the accomplishment of a designated task according to pre-established or given specifications. The learner may be required to provide a verbal description to accompany the performance.

Examples:

1. Objective: Given a subject in any prostrate position, the learner will demonstrate the proper method for administering mouth-to-mouth resuscitation.
   Test Item: Demonstrate on your partner the proper procedure for administering mouth-to-mouth resuscitation.
2. Objective: Given a ruler and compass, the learner will demonstrate and describe the procedure for constructing an isosceles triangle.
   Test Item: Show the correct procedure to use in drawing an isosceles triangle.

   Equivalent terms and phrases: show your work, show the procedure, perform an experiment, perform the steps.

   (The basis of evaluation in "demonstrating" is the set of procedures that the learner performs in order to produce a product, select a response, or complete a particular act, and not the selected response or constructed product itself.)

Although there are few learner behaviors on cognitive tasks that cannot be classified under one of the six terms in the above list, there are certain occasions in constructing statements of objectives when it is more convenient to use other terms. Some verbs that may be employed in writing objectives for specific subject-matter areas have meanings that are so precise that they require no further explication. Such words as spell, subtract, read and alphabetize should be used in preparing statements of objectives whenever they are appropriate. The above list, however, names and defines behaviors that are common to many curriculum areas and provides a framework for constructing precise objectives for most instructional tasks.

The objective used as the second example for each of the six performance terms in the list above can serve to illustrate a procedure for using the list. In classroom instruction the teacher sometimes fails to consider the development of appropriate learner behaviors toward given content because these potential behaviors do not occur to him. For example, in instruction about such diverse content as various principles of learning in an educational psychology course or isosceles triangles in an elementary school mathematics course, the learner is frequently required only to provide a verbal definition of the particular phenomenon (e.g., an isosceles triangle is a triangle having two equal sides), rather than being expected to acquire the ability to identify new examples or produce his own examples of it. In using the six performance terms to identify appropriate learner behaviors, the teacher or curriculum designer starts with a given bit of content (in this case, isosceles triangles) to be treated in his course. He then uses the

performance terms as a checklist to identify and state those behaviors that he wants his students to acquire toward that content. To the extent that the curriculum designer and teacher are able to identify and develop desirable en route and terminal learner behaviors that they would otherwise overlook for any given content, the use of the six performance terms or a similar behavioral classification scheme as a checklist is an important aid to curriculum planning and development.

It can be seen from the six examples dealing with isosceles triangles in the above list that for some content it is possible through a checklist approach to identify appropriate learner outcomes toward that content within all six classes. However, with other units of content it is not always possible to identify appropriate learner performance within each of the six classes. In addition, the teacher or curriculum designer will sometimes determine that it is not desirable (for reasons of time, student ability, curriculum scope and sequence, etc.) to develop in his course all behaviors that he is able to identify for specific units of course content.

The reader who has completed Mager's (1962) excellent book, Preparing Instructional Objectives, will recognize the first exemplar objective in the above list for the term "identify" (given a list of instructional objectives, the learner will identify those that are stated in terms of learner behavior) as a terminal objective for the Mager book. Instructional objectives, the content covered in Mager's book, can be used to further explicate the use of the six performance terms. Suppose that the instructor in a course in curriculum and instruction decides that he wants his students to learn about behavioral objectives so that they will use them in teaching. Mager's book, which teaches the successful learner to identify behavioral objectives written by someone else, is frequently used for this purpose. Yet, if the instructor uses the six performance terms in checklist fashion, he will undoubtedly determine that it is important for his students to be able to construct (i.e., write) their own behavioral objectives as well as to identify given examples and non-examples of behavioral objectives written by someone else. He would therefore include for his course an objective such as "The learner will construct statements of behavioral objectives in his own (or any specified) subject-matter area". If the instructor used Mager's book to teach students to identify examples of behavioral objectives, he would also need to provide instruction, practice and evaluation on the writing of objectives. By indicating this additional important outcome, the checklist would assist the instructor in formulating the objectives and instruction for his course.

Although the six performance terms are not intended to constitute a hierarchy of behaviors that are sequenced in a fixed order across all content, for given tasks it is possible to suggest seemingly optimal sequences for developing two or more of the listed behaviors. For example, it is likely that the development of constructing and demonstrating behaviors is facilitated by learner acquisition of verbal chains that mediate the desired performances. Thus, instruction would be sequenced so that these chains, subsumed under the terms "describe" and "order",

are acquired before practice is provided on the constructing and demonstrating tasks. In turn, efficient use of a verbal chain or description to mediate constructing or demonstrating behavior depends upon the learner's ability to identify examples of the concepts contained in the description. This suggests the importance of developing the learner's ability to identify examples of these concepts prior to his acquisition of the mediating description. Again, for concept-attainment tasks, discovery-learning devotees presumably would support an approach under which the learner initially identifies or names instances of the concept, then arrives at a description of its attributes, and later, if both feasible and desirable, constructs examples of it. Their deductive-oriented counterparts, on the other hand, would be more inclined to advocate a procedure in which the learner first describes the attributes of the concept, then identifies and/or names examples of it, and finally constructs examples.

While a logical analysis supports the use of the sequencing procedures suggested above, validation of their efficiency awaits empirical evidence on their use in school curricula. The six performance terms provide one standard framework for describing tasks in a way that testable hypotheses can be formulated on the most efficient order of instructional tasks within a curriculum and on sequencing strategies across curricula. The design of curricula that are sequenced on the basis of empirical evidence must involve the formulation and testing of such hypotheses.

OBJECTIVES AND EVALUATION

Curriculum experts have emphasized the importance of precise instructional objectives for two primary purposes: planning instruction and assessing its effects. Instructional planning and the assessment of learner performance are so closely interrelated that one can hardly be considered in detail independently of the other. Good instructional planning is based upon an assessment of the skills possessed by the intended student population, and the evaluation of instruction obviously must be based upon measurement of its outcomes. The information that learners have or have not acquired given behaviors serves no useful educational purpose in itself. However, this information can be used to make sound decisions about instructional treatments. The use of instructional objectives in evaluation can lead to educational improvement by resulting in the development and adoption of more effective curricula and by revealing the learning deficiencies of individual students and indicating appropriate treatments to overcome them.

FORMATIVE EVALUATION

Scriven (1967) has introduced the term "formative evaluation" to describe the evaluation of educational programs that are still in some stage of development. This contrasts with summative evaluation, or the evaluation of programs in a finally developed form. A major concern in

formative evaluation is the identification of materials and procedures that will increase the effectiveness of the program being evaluated. Summative evaluation more often takes the form of post-development comparison between two or more existing programs, with little or no attention paid to procedures for improving either program. The product of formative evaluation activities is expected to be an improved instructional program, while the product of summative evaluation is normally a set of descriptive statements about the efficacy of a single program or the relative merits of two or more programs.

Formative evaluation procedures can be employed both in the development of new curricula or curriculum materials and in the teacher's evaluation of his own classroom instruction in given courses, since in each case revisions in instructional materials and methods can be accomplished following the evaluation. However, the particular procedures used in formative evaluation activities will vary considerably with the instructional program being evaluated. Ideally, extensive evaluation will be conducted with any new curriculum that is to be used by a large number of students. Initial procedures with early drafts of prototype materials might involve their tryout with individual learners and small groups, followed by revisions based upon learner performance. Field testing of the prepared materials should be conducted in a number of classrooms, with frequent observations to determine whether or not the program is actually presented as expected. Individual testing of learners participating in the field tryouts will be employed when, as in the case of a primary-grade reading program, it is the only appropriate method for measuring learner performance on certain important desired outcomes of the program. More measurement may be included than would normally occur in regular classroom instruction and evaluation, so that performance can be measured on each objective and weaknesses in the program can be identified and corrected. Because of their potential effect in subsequently leading to improved performance by large numbers of learners, many procedures that are too costly or time consuming for teacher use in the evaluation of ongoing classroom instruction will be employed to evaluate and improve new curricular programs.

Assessment based upon instructional objectives is a crucial part of well-designed formative evaluation. A course of instruction or new curriculum is established so that learners will acquire certain capabilities. The presence or absence of these capabilities in a learner must be judged on the basis of whether or not he manifests them in appropriate situations. Before the instructional activities for a course are identified, the desired behavioral outcomes must be clearly stated so that activities designed to produce them can be selected. The most important purpose of formative evaluation in a carefully planned instructional program is to indicate the desired outcomes that learners do not acquire at an acceptable performance level. The teacher or curriculum developer (they may be the same person in cases when the teacher is evaluating his own instructional unit) is then able to design, implement and evaluate potential improvements in the instruction related to these objectives.

In formative evaluation endeavors there are several possible approaches for identifying the objectives for which instruction should be revised. One method does not entail the specification of performance standards for each stated objective prior to instruction. Instead, post-instructional learner performance is measured and a score, normally in percentage form, is derived to summarize learner performance for each objective. An analysis that takes into account the judged importance of each objective and the overall performance score on each is conducted with these data to identify the objectives for which learner performance is unsatisfactory. Revisions designed to result in improved performance by new learners are then made in the instructional materials and procedures for these objectives.

Other formative evaluation procedures for identifying the objectives for which the instruction should be revised involve the specification prior to instruction of minimal acceptable performance standards for each objective. Specification of minimal performance levels prior to instruction seems particularly appropriate when the objectives represent behaviors whose mastery is essential to the subsequent attainment of other important objectives. For instance, mastery of the objectives of primary-grade units on basic addition and subtraction facts is necessary for future success on many arithmetic tasks. It is probably more important to set minimal standards prior to instruction and to conduct successive trial and revision cycles until new learners attain these standards with instructional units of this type than it is with, say, a twelfth-grade course in modern English literature.

Performance standards may be set for an objective prior to instruction by designating a single minimum acceptable raw score or percentage score for all learner responses to all criterion items for that objective, or by specifying both student and class minimal levels as advocated by Baker (1966). The student performance level for an objective refers to the desired minimal score that students are expected to attain on that objective following instruction. The class level refers to the percentage of students in the class or other target group who will attain the stated student minimal level if the instruction is successful. For example, in the following objectives from a beginning reading program, the class minimal level is underlined once and the student minimal level twice:

1. Given each of the 10 printed letters taught in the initial unit, 90 per cent of the learners will state the most common sound "made by" at least 9 of the 10. (The letters are l, m, n, r, s, a, e, i, o, u.)

2. Given previously unencountered, regularly spelled, one-syllable words composed exclusively from the 10 printed letters, 70 per cent of the learners will correctly read 75 per cent of the words.

The objectives above also illustrate the necessity of establishing high performance levels on certain tasks. Mastery of the first objective is essential to success on the second objective. Therefore, it is very important that high performance levels be established for the first

objective. If formative evaluation trials reveal sub-standard learner performance, successive revisions of the relevant instruction should be made until new learners attain the established standards. For the second objective, on the other hand, there is evidence (Silberman, et al., 1964) that the difficulty level for young children is such that it is more realistic to set comparatively low performance levels. Higher standards may ultimately be established in later grades for an expanded form of the same objective (e.g., one that includes all 26 letters and multi-syllabic words).

A warning should be sounded here against the frequent derogation by educators of so-called "low-level tasks", as exemplified by the first reading objective above, in favor of objectives that are higher level in the sense of the presumed mental operations required for their attainment. Failure to acquire key en route behaviors will inevitably lead to failure to acquire desired "higher level" terminal behaviors. Yet, some educators expect students to somehow ascend from the ground floor to an upper cognitive level of their educational edifice without ever climbing the stairs or even riding the elevator past the lower levels. An incident that occurred during a recent national meeting of members of a prominent educational research association illustrates the case in point. A distinguished curriculum expert who played a very active role in the development of one of the new science curricula made the curious and somewhat incredible remark that, quite frankly, she did not "give a damn about whether or not students acquire behaviors below level 3 of the (Bloom) taxonomy." In light of her remark, it is fortunate that several studies with the taxonomy reveal that educators' verbal statements purportedly indicating the types of learner performance that they consider important do not accurately reflect the tasks which they themselves emphasize in class.

In addition to the judged importance of the behaviors as en route or terminal instructional outcomes for the target group, other factors that should be considered in establishing performance standards include the difficulty level of each designated behavior and the amount of instructional time available. Data obtained from pre-instructional and post-instructional testing can provide an empirical basis for judging the importance of en route behaviors and setting performance standards that are realistic for the difficulty level of the task. Successive trial and revision cycles may eventually reveal the need to revise some established performance standards downward. However, initial evaluation and revision efforts should concentrate upon improving learner performance on those activities for which the established post-instructional standards are not attained.

Irrespective of whether performance standards set prior to instruction contain both student and class minimal levels or a single score based upon all learner responses for that objective, essentially the same procedure is employed following instruction to identify objectives for which the instruction should be revised. Criterion items for each objective

are administered to all students or to a random sample of students from the target group, and actual post-instructional performance levels are computed to show either the student and class performance or the single score for each objective. The relevant instruction is then revised for each objective on which post-instructional learner performance does not equal or surpass the minimal performance standards established prior to instruction.

Once the curriculum developer or teacher has identified the instruction related to each objective on which learners did not attain the established standards, he is faced with the task of determining instructional refinements that will improve learner performance on the objectives. There are certain procedures that will frequently enable an individual to identify refinements in the instruction for a given objective. One procedure simply involves a logical analysis of the relevant instruction to determine the need for increased simplicity or more redundancy in the information presented, more examples to explicate the information, and/or additional practice and knowledge of results on that particular task. Of course, empirical evidence on the accuracy of this logical analysis is collected in subsequent tryouts of the instructional sequence. Analysis of the incorrect responses of learners on both practice and criterion items related to an objective may also enable the curriculum developer to identify appropriate refinements by revealing certain types of errors consistently made by learners. For multiple-choice items this is especially true if care is taken initially to select distractors that represent classes of errors common to the specific task, but analysis of incorrect responses may be equally effective on other types of criterion ems. Schutz, Baker and Sullivan (1967) report the use of structured individual interview techniques with a small number of learners at given points during an instructional unit as another technique for identifying instructional improvements for particular low-performance objectives. Teachers who have used the instructional materials in the classroom may suggest procedures or additional materials that will improve learner performance on specified objectives, and observation of the relevant instruction in the classroom will also frequently indicate effective refinements.

In attempting to identify potential improvements in an instructional program, it does the curriculum developer little good to know that students' pre-instructional and post-instructional average grade-placement scores were 2.2 and 3.1 respectively, or that their mean score on the criterion test was 73 per cent. However, it is extremely important for instructional improvement purposes that he know which desired outcomes were achieved by the students and which were not. Formative evaluation procedures based upon assessment of learner performance on the program's instructional objectives provide the only systematic method for obtaining this information. The proper use of such procedures should frequently result in significant improvements in school curricula.

SUMMATIVE EVALUATION

Summative evaluation differs from formative evaluation in that summative efforts are not designed to produce improvements in a given set of materials or procedures. Summative evaluation studies are conducted either for the purpose of comparing the efficacy of two or more programs or in order to determine the effects of a single program. For the latter purpose, evaluation is occasionally based upon learner performance on pre-specified instructional objectives, but the normal procedure in summative studies is to measure learner performance with commercially published standardized tests.

The difficulties associated with the use of standardized tests for evaluating educational programs are well known. Faced with no-significant-difference results when comparing his program with traditional ones, the developer of a new curriculum indignantly, and often justifiably, denounces the validity for his curriculum of the very test that he selected for evaluation purposes. The standardized test, he claims, was constructed to measure the content of textbooks and curriculum guides in use prior to the development of his new curriculum. In actual fact, standardized tests are not constructed to sample very specifically the content of any single instructional program, although quite frequently they sample less adequately the specific content and behaviors featured in new curricula.

Certainly there is a serious need for improved summative evaluation procedures in selecting the best curriculum for school use from among competing programs. Because objectives vary to some extent from program to program, it will be difficult to ever develop a method for comparing programs that will be completely satisfactory to the sophisticated researcher or even minimally acceptable to the curriculum developer whose program fails to show superior results. However, in comparison to their unique aspects, most programs in the same subject area for given grade levels have a large number of common features. One approach that may have merit in evaluating competing programs is to construct a test or tests that sample both the common objectives of the programs and the objectives that are unique to each program. Learner performance can then be compared on specific objectives that are common to both programs, as well as on the objectives that are unique to each. As in other measurement procedures based upon stated or inferred behavioral objectives, the evaluator would use criterion-referenced, as opposed to norm-referenced, test construction procedures (see Glaser, 1963). Like any other method for evaluating competing programs, this approach has certain limitations and would require value judgments by subject-matter experts about the relative merit of the objectives or outcomes that are unique to each program. Nevertheless, in cases where it is feasible, it would appear to be superior to the use of a standardized test that is not keyed to the specific objectives of either program that is being evaluated.

The observed capabilities that learners acquire from a school curriculum should be the matter of utmost concern in its evaluation. Normally these

capabilities can be accurately assessed by measuring learner performance on the instructional objectives of the curriculum. Also of importance in evaluation are teacher and pupil attitudes toward the program, since these attitudes may be useful for revision purposes or they may influence the adoption and effective use of a program. Other criteria which have much less merit have been suggested for evaluation purposes. For example, Stake (1966) reports that the systematically gathered judgments of "groups having opinions on education", such as parents and spokesmen for society at large, are essential data for evaluation of school curricula. These groups simply are not competent evaluators of school programs. Their opinions often are so much a function of the school's public relations efforts that a mere popular title and glowing description of a program will greatly influence its acceptance. Also, it is questionable whether significant complex educational outcomes that are incidental to the objectives of a well-planned instructional program occur as frequently as suggested by Eisner (1966). It is very doubtful that these initially unintended and indeterminable outcomes are worthy of intensive evaluation. The important desirable outcomes of instruction are just not that difficult to determine in advance, and the greatest potential pay-off in evaluation efforts lies in the identification and subsequent adoption of procedures that produce these identifiable outcomes.

PRESCRIBING INDIVIDUALIZED INSTRUCTION

In recent years considerable attention has been given to the utility of behavioral objectives for the purpose of assigning different instructional treatments to individuals within a course. Grouping within and among classes on the basis of individual differences in achievement has, of course, long been a standard practice. The normal bases for selecting the members of each group have been such general factors as teacher recommendations and standardized aptitude and achievement test scores. The assignment of instructional treatments to individuals based upon their performance on specified objectives represents an attempt to achieve greater precision and effectiveness in diagnosing and meeting individual needs.

Depending upon whether instruction is individually paced or group-paced, somewhat different procedures are used for assigning treatments to students based upon their performance on instructional objectives. Individually paced instruction obviously requires self-study materials. In an individually paced course the student is initially placed at an entering level in an instructional sequence. This initial level is determined by one or more examinations that measure his performance on the requisite entering behaviors and the course objectives. He starts his instruction on the material related to the first task that he has not mastered in the instructional sequence, and he works independently with that material until his post-instructional test performance reveals that he has mastered that task at a pre-specified performance level. Concurrently, the student's performance on upcoming objectives in the instructional sequence is measured so that he is able to bypass instruction on the objectives that he has already mastered. The importance of maintaining control over his progression from task to task is apparent when

one considers that mastery of certain en route tasks will often be essential to acceptable performance on later tasks. Glaser (1966) discusses this issue in greater detail and presents a description of an individually paced program in school use.

Because of factors such as the limited reading repertoire of young children, instruction cannot always be individually paced for all classes even if it is deemed desirable. Objectives-keyed mastery tests can also be used to diagnose individual weaknesses and prescribe appropriate differentiated treatments to individuals within group-paced classroom instruction, irrespective of whether the group consists of the entire class or a more homogeneous group within the class. Work with systems of this type is currently in progress at the Southwest Regional Laboratory. Under one approach a test that measures performance on three or four recent instructional objectives is included in the lesson materials for every Friday. Learner performance on the test is analyzed by computer, and a print-out is generated that lists for each objective the learners who fail to attain the specified criterion level. Each learner is then assigned a prepared instructional exercise for that objective. He completes the exercise in an activity period during the week following the Friday test to which the instructional exercise relates. Each learner proceeds with his group during the greater part of the week, but as a built-in feature of the program he also receives instruction and practice designed to remedy his particular identified deficiencies before they become serious handicaps to subsequent learning.

THE FUTURE FOR OBJECTIVES-BASED EVALUATION

It has been suggested in this paper that both the effectiveness of new curricula and the achievement of individual learners can be improved by the use of certain evaluation procedures based upon precise instructional objectives. However, the author also warned that the presence or absence of a set of one's own cherished procedures in a program is not a proper basis for evaluating that program. Instead, the procedures and program must be evaluated on the basis of their effects on learner performance.

There is clear need for further refinement and study of the efficacy of evaluation procedures such as those suggested in this paper. Remarkably few major projects in the development of new curricula or in the improvement of classroom instruction have attempted to employ systematic assessment of learner performance on precise instructional objectives as a means for instructional improvement. Consequently, objectives-based techniques for conducting formative evaluation efforts with new curricula and for individualizing classroom instruction are in the early formative evaluation stages themselves. Their refinement is contingent upon evaluation of their use in the develop of instructional programs and in ongoing classroom instruction. ch procedures have been sufficiently refined, there will still need for summative evaluation to compare

programs that incorporate them with those that do not. Like the methods and materials in any instructional program, the value of these evaluation procedures ultimately must be judged by their effects on student performance.

Research and evaluation efforts may eventually produce convincing evidence that improved learner performance results from the use of precise instructional objectives both to evaluate and revise instruction and to diagnose and remedy individual weaknesses. Nevertheless, there is little reason to expect that teachers will rush to adopt these procedures. The evaluation of instruction and of individual learners, and the subsequent modification of instruction based upon learner performance on pre-specified objectives is a difficult task requiring considerable time and effort. Individual teachers cannot be expected to accomplish this task exclusively on their own efforts. It is much easier for teachers to identify classroom activities simply by asking the aforementioned "What am I going to do?" question. Teachers learn in much the same manner as their students. If we really expect them to change their behavior so that they employ effective new instructional procedures in the classroom, we must set up appropriate conditions to establish and maintain the new behavior. Just as telling students how to perform a task is not a sufficient condition to ensure mastery of the task, merely telling teachers how to improve their instruction or presenting them with impressive research data is not likely to lead to improved instruction.

Two conditions appear to be essential if teachers are to use pre-specified instructional objectives to evaluate and revise instruction and to prescribe treatments for specific deficiencies of individual learners. First, procedures for accomplishing the task must be simplified so that the teacher is not required to design the entire evaluation system himself. The demands on teachers are such that they do not have time to design and develop this type of system for their courses, nor does their training typically prepare them to do so. Materials that will greatly reduce the demands on the teacher in an objectives-based evaluation system can and should be incorporated directly into prepared instructional programs. The required teacher time and effort will not be inordinate if he is given as a part of the instructional program the requisite stated objectives, mastery and criterion tests keyed to the objectives, remedial exercises, routine test-scoring and print-out service, and directions for using the materials. The teacher who uses such a system still maintains control over the instructional decision-making process (i.e., he chooses the objectives, instructional activities, remedial exercises, etc.), but he is provided with an array of materials that enables him to diagnose individual and group deficiencies and prescribe remedial activities with much greater precision than his time otherwise permits.

The second necessary set of events relates to the training and classroom conditions for teacher acquisition and maintenance of these procedures. Teachers will require both instruction and considerable practice in the use of the procedures. In addition, they must receive reinforcement for using the procedures to improve learner performance in the regular classroom

setting if they are to continue to use them consistently. Current methods for evaluating and rewarding teachers place little or no emphasis on the acquisition of desired learning outcomes by their students (see McNeil, 1966).

The development and proper classroom installation of effective objectives-based instructional systems will require the combined efforts of learning specialists, curriculum and subject-matter experts, and school administrators and teachers. These efforts can best be accomplished under the direction of agencies that are staffed by able professional scholars and talented school personnel committed to the systematic development and evaluation of new instructional methods and materials. Such efforts, if successful, will have the much-needed consequence of significantly improving student achievement in our schools.

REFERENCES

American Association for the Advancement of Science. Commission on Science Education Newsletter, 1965, I, 2-4.

Anderson, R. C. Education psychology. Annual Review of Psychology, 1967, 18, 129-164.

Arnstine, D. G. The language and values of programmed instruction. The Educational Forum. 1964, 28, 337-45.

Baker, Eva. Establishing performance standards. Los Angeles: Vimcet Associates, 1966.

Bialek, H. M. A measure of teachers' perceptions of Bloom's educational objectives. Paper read at the annual meeting of the American Educational Research Association, New York, February, 1967.

Bloom, B. S., (Ed.) A taxonomy of educational objectives: Handbook I, the cognitive domain. New York: Longmans, Green, 1956.

Cox, R. C., & Unks, Nancy. A selected and annotated bibliography of studies concerning the taxonomy of educational objectives: Cognitive domain. University of Pittsburgh: Learning Research and Development Center, 1967.

Davis, O. L., & Tinsley, Drew. Cognitive objectives revealed by classroom questions asked by social studies student teachers. Paper read at the annual meeting of the American Educational Research Association, New York, 1967.

Eisner, E. W. Educational objectives: Help or hindrance. Paper read at the annual meeting of the American Educational Research Association, Chicago, February, 1966.

Gagne, R. M. The conditions of learning. New York: Holt, Rineheart & Winston, 1965.

Gerlach, V. S., & Sullivan, H. J. Constructing statements of outcomes. Inglewood, Calif: Southwest Regional Laboratory for Educational Research and Development, 1967.

Glaser, R. L. Instructional technology and the measurement of learning outcomes: Some questions. American Psychologist, 1963, 18, 519-521.

Glaser, R. L. The program for individually prescribed instruction. Paper read at the annual meeting of the American Educational Research Association, Chicago, February, 1966.

Jackson, P. W. The way teaching is. Washington, D.C. Association for Supervision and Curriculum Development, 1966.

Krathwohl, D. R. The taxonomy of educational objectives -- Its use in curriculum building. In C. M. Lindvall (Ed.), Defining educational objectives. Pittsburgh: University of Pittsburgh Press, 1964.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. Taxonomy of educational objectives. Handbook II: Affective domain. New York: McKay, 1964.

Kropp, R. P. & Stoker, H. W. The construction and validation of tests of the cognitive processes as described in the taxonomy of educational objectives. Cooperative Research Project No. 2117, Florida State University: Institute of Human Learning, 1966.

Lawrence, C. D. Analysis of teacher tests in social studies according to the taxonomy of educational objectives. Claremont, Calif: Claremont Colleges, Honnold Library, 1963.

Macdonald, J. B. Myths about instruction. Educational Leadership. 22, 1965, 571-576.

Mager, R. F. Preparing instructional objectives. Palo Alto: Fearon Publishers, 1962.

McGuire, Christine. Research in the process approach to the construction and to analysis of medical examinations. Paper read at annual meeting of National Council on Measurement in Education, Chicago, February, 1963.

McNeil, J. D. Antidote to a school scandal. Educational Forum, 31, 1966, 69-77.

Popham, W. J. Objectives and instruction. University of California, Los Angeles, 1968. (Mineographed)

Scannell, D. P., & Stellwagon, W. R. Teaching and testing for degrees of understanding. California Journal of Instructional Improvement, 3, 1960, 1.

Schutz, R. E., Baker, R. L., & Sullivan, H. J. A guide to the development of instructional materials. Technical Report, Aerospace Medical Research Laboratory, Wright-Patterson AF Base, Ohio, 1967.

Scriven, M. The methodology of evaluation. In B. O. Smith (Ed.), Perspectives of curriculum evaluation. Chicago: Rand McNally, 1967.

Silberman, H., Coulson, J., Melaragno, R., & Newmark, G. Use of exploratory research and individual tutoring techniques for the development of programming methods and theory. Title VII Research Report, Santa Monica: System Development Corporation, 1964.

Stake, R. E. The countenance of educational evaluation. University of Illinois, Center for Instructional Research and Curriculum Evaluation, 1966. (Mimeographed)

Stanley, J. C. (Ed.) Improving experimental design and statistical analysis. Chicago: Rand McNally, 1967.

Stanley, J. C., & Bolton, D. L. A review of Bloom's "Taxonomy of educational objectives" and J. R. Gerberich's "Specimen objective test items, a guide to achievement test construction". Educational and Psychological Measurement, 17, 1957, 631-634.

Sullivan, H., Baker, R. L., & Schutz, R. E. Classroom episodes for teaching psychological principles and concepts of learning. Arizona State University, Classroom Learning Laboratory, 1966.

Wolf, R. Review of "The construction and validation of tests of the cognitive processes as described in the taxonomy of educational objectives". Educational and Psychological Measurement, 27, 1967, 542-548.